Supplementary material A class of implicit-explicit two-step Runge-Kutta methods

Evgeniy Zharovsky, Adrian Sandu, and Hong Zhang

Appendix A. Prothero-Robinson convergence.

We consider the Prothero-Robinson (PR) [18] test problem written as a split system (1.1)

$$y' = \underbrace{\mu(y - \phi(t))}_{q(y)} + \underbrace{\phi'(t)}_{f(y)}, \quad \mu < 0, \quad y(0) = \phi(0),$$
 (A.1)

where the exact solution is $y(t) = \phi(t)$. A numerical method is said to be PRconvergent with order p if its application to (A.1) gives a solution whose the global error decreases as $\mathcal{O}(h^p)$ for $h \to 0$ and $h\mu \to -\infty$.

THEOREM A.1. Consider the IMEX TSRK method (3.1) of order p, stage order q for the explicit part, and stage order \hat{q} for the implicit part. Assume that the implicit part is linearly stable, and that the spectral radius of the implicit stability matrix (4.3) is bounded uniformly in the infinite "region of interest" (5.1)

$$\rho\left(\widehat{\mathbf{M}}(z)\right) \le \rho_0 < 1, \quad \forall \ z \in \mathcal{R}.$$

Then the IMEX method (3.1) is PR-convergent with order min(p,q).

In particular if the explicit stage order is q = p, then the PR order of convergence is p. It is convenient to construct IMEX TSRK methods (3.1) with explicit stage order q = p, even if $\hat{q} = p - 1$, as such methods do not suffer from order reduction on the PR problem.

Proof. Let

$$\phi^{[n]} = \phi(t_{n-1} + \mathbf{c} h) = [\phi(t_{n-1} + c_1 h), \dots, \phi(t_{n-1} + c_s h)]^T$$
.

The method (3.1) applied to (A.1) reads:

$$Y^{[n]} = (\mathbf{e} - \mathbf{u}) \ y_{n-1} + \mathbf{u} \ y_{n-2}$$

$$+ h \mathbf{A} \ \phi'^{[n]} + h \mathbf{B} \ \phi'^{[n-1]}$$

$$+ h \mu \widehat{\mathbf{A}} \left(Y^{[n]} - \phi^{[n]} \right) + h \mu \widehat{\mathbf{B}} \left(Y^{[n-1]} - \phi^{[n-1]} \right) ,$$

$$y_n = (1 - \vartheta) \ y_{n-1} + \vartheta \ y_{n-2}$$

$$+ h \mathbf{v}^T \ \phi'^{[n]} + h \mathbf{w}^T \ \phi'^{[n-1]}$$

$$+ h \mu \mathbf{v}^T \left(Y^{[n]} - \phi^{[n]} \right) + h \mu \mathbf{w}^T \left(Y^{[n-1]} - \phi^{[n-1]} \right) .$$
(A.2a)

Consider the global errors

$$e_n = y_n - \phi(t_n) , \quad E^{[n]} = Y^{[n]} - \phi^{[n]} .$$

Write the stage equation (A.2a) in terms of the exact solution and global errors

$$\begin{split} E^{[n]} &= -\phi \, (t_{n-1} + \mathbf{c} \, h) + (\mathbf{e} - \mathbf{u}) \, \, \phi(t_{n-1}) + \mathbf{u} \, \phi(t_{n-2}) \\ &\quad + (\mathbf{e} - \mathbf{u}) \, \, e_{n-1} + \mathbf{u} \, e_{n-2} \\ &\quad + h \, \mathbf{A} \, \phi' \, (t_{n-1} + \mathbf{c} \, h) + h \, \mathbf{B} \phi' \, (t_{n-2} + \mathbf{c} \, h) \\ &\quad + h \, \mu \, \widehat{\mathbf{A}} \, E^{[n]} + h \, \mu \, \widehat{\mathbf{B}} \, E^{[n-1]} \, , \end{split}$$

to obtain

$$(\mathbf{I} - h \,\mu \,\widehat{\mathbf{A}}) E^{[n]} = (\mathbf{e} - \mathbf{u}) e_{n-1} + \mathbf{u} e_{n-2} + h \,\mu \,\widehat{\mathbf{B}} E^{[n-1]}$$

$$- (\phi (t_{n-1} + \mathbf{c} h) - \mathbf{e} \phi(t_{n-1})) + \mathbf{u} (\phi(t_{n-2}) - \phi(t_{n-1}))$$

$$+ h \,\mathbf{A} \,\phi' (t_{n-1} + \mathbf{c} h) + h \,\mathbf{B} \phi' (t_{n-2} + \mathbf{c} h) .$$

$$(A.3)$$

The exact solution is expanded in Taylor series about t_{n-1} :

$$\phi(t_{n-1} + \mathbf{c} h) - \mathbf{e} \phi(t_{n-1}) = \sum_{k=1}^{\infty} \frac{h^k \mathbf{c}^k}{k!} \phi^{(k)}(t_{n-1}),$$

$$\phi(t_{n-2}) - \phi(t_{n-1}) = \sum_{k=1}^{\infty} \frac{(-1)^k h^k}{k!} \phi^{(k)}(t_{n-1}),$$

$$h \phi'(t_{n-1} + \mathbf{c} h) = \sum_{k=1}^{\infty} \frac{k h^k \mathbf{c}^{k-1}}{k!} \phi^{(k)}(t_{n-1}),$$

$$h \phi'(t_{n-2} + \mathbf{c} h) = \sum_{k=1}^{\infty} \frac{k h^k (\mathbf{c} - \mathbf{e})^{k-1}}{k!} \phi^{(k)}(t_{n-1}).$$

Inserting the above Taylor expansions in (A.3) leads to

$$\begin{split} \left(\mathbf{I} - h \, \mu \, \widehat{\mathbf{A}} \right) \, E^{[n]} &= (\mathbf{e} - \mathbf{u}) \, e_{n-1} + \mathbf{u} \, e_{n-2} + h \, \mu \, \widehat{\mathbf{B}} \, E^{[n-1]} \\ &\quad + \sum_{k=1}^{\infty} \left(-\mathbf{c}^k + (-1)^k \, \mathbf{u} + k \, \mathbf{A} \, \mathbf{c}^{k-1} + k \, \mathbf{B} \, (\mathbf{e} - \mathbf{c})^{k-1} \right) \, \frac{h^k}{k!} \phi^{(k)}(t_{n-1}) \\ &= (\mathbf{e} - \mathbf{u}) \, e_{n-1} + \mathbf{u} \, e_{n-2} + h \, \mu \, \widehat{\mathbf{B}} \, E^{[n-1]} + \mathcal{O} \left(h^{q+1} \right) \end{split}$$

where q is the stage order of the explicit method. The last equality follows from the stage order conditions (2.3). Let $z = h\mu$. We have:

$$E^{[n]} = \widehat{\mathbf{S}}(z) \ (\mathbf{e} - \mathbf{u}) \ e_{n-1} + \widehat{\mathbf{S}}(z) \mathbf{u} \ e_{n-2} + z \ \widehat{\mathbf{S}}(z) \ \widehat{\mathbf{B}} \ E^{[n-1]} + \widehat{\mathbf{S}}(z) \ \mathcal{O} \left(h^{q+1} \right) , (A.4)$$

$$\widehat{\mathbf{S}}(z) = \left(\mathbf{I} - z \ \widehat{\mathbf{A}} \right)^{-1} .$$

From (A.4) it follows that

$$z \mathbf{v}^{T} E^{[n]} = z \mathbf{v}^{T} \widehat{\mathbf{S}}(z) (\mathbf{e} - \mathbf{u}) e_{n-1} + z \mathbf{v}^{T} \widehat{\mathbf{S}}(z) \mathbf{u} e_{n-2}$$

$$+ z^{2} \mathbf{v}^{T} \widehat{\mathbf{S}}(z) \widehat{\mathbf{B}} E^{[n-1]} + z \widehat{\mathbf{S}}(z) \mathcal{O} (h^{q+1}) .$$
(A.5)

Note that h and $h\mu$ are allowed to vary independently, and therefore the order of the asymptotic term does not change upon multiplication by $z = h\mu$.

Similarly, write the solution equation (A.2b) in terms of the exact solution and global errors:

$$e_{n} = -\phi(t_{n}) + (1 - \vartheta) \phi(t_{n-1}) + \vartheta \phi(t_{n-2})$$

$$+ (1 - \vartheta) e_{n-1} + \vartheta e_{n-2}$$

$$+ h \mathbf{v}^{T} \phi' (t_{n-1} + \mathbf{c} h) + h \mathbf{w}^{T} \phi' (t_{n-2} + \mathbf{c} h)$$

$$+ h \mu \mathbf{v}^{T} E^{[n]} + h \mu \mathbf{w}^{T} E^{[n-1]}.$$

After rearranging the expression, and expanding the exact solution in Taylor series about t_{n-1} , we obtain

$$e_{n} = (1 - \vartheta) e_{n-1} + \vartheta e_{n-2} + h \mu \mathbf{v}^{T} E^{[n]} + h \mu \mathbf{w}^{T} E^{[n-1]}$$

$$+ \sum_{k=1}^{\infty} \left(-1 + (-1)^{k} \vartheta + k \mathbf{v}^{T} \mathbf{c}^{k-1} + k \mathbf{w}^{T} (\mathbf{c} - \mathbf{e})^{k-1} \right) \frac{h^{k}}{k!} \phi^{(k)}(t_{n-1})$$

$$= (1 - \vartheta) e_{n-1} + \vartheta e_{n-2} + h \mu \mathbf{v}^{T} E^{[n]} + h \mu \mathbf{w}^{T} E^{[n-1]} + \mathcal{O} (h^{p+1}) . \quad (A.6)$$

The last equality follows from the order conditions (2.5).

The following error recurrence is obtained by combining (A.4), (A.5), and (A.6)

$$\begin{bmatrix} e_n \\ e_{n-1} \\ E^{[n]} \end{bmatrix} = \widetilde{\mathbf{M}}(h\mu) \cdot \begin{bmatrix} e_{n-1} \\ e_{n-2} \\ E^{[n-1]} \end{bmatrix} + \mathcal{O}\left(h^{\min(p+1,q+1)}\right). \tag{A.7}$$

Assume a one-step, order p method is used to initialize both the step and the stage solutions of the TSRK method [16, Section 6.2]. The error starting values are $e_0 = 0$, $e_1 = \mathcal{O}(h^p)$, and $E^{[1]} = \mathcal{O}(h^p)$. The error amplification matrix

$$\widetilde{\mathbf{M}}(z) = \begin{bmatrix} 1 - \vartheta + z \, \mathbf{v}^T \, \widehat{\mathbf{S}}(z) \, \left(\mathbf{e} - \mathbf{u} \right) & \vartheta + z \, \mathbf{v}^T \, \widehat{\mathbf{S}}(z) \, \mathbf{u} & z \, \left(\mathbf{w}^T + z \, \mathbf{v}^T \, \widehat{\mathbf{S}}(z) \, \widehat{\mathbf{B}} \right) \\ 1 & 0 & 0 \\ \widehat{\mathbf{S}}(z) \, \left(\mathbf{e} - \mathbf{u} \right) & \widehat{\mathbf{S}}(z) \, \mathbf{u} & z \, \widehat{\mathbf{S}}(z) \, \widehat{\mathbf{B}} \end{bmatrix} ,$$

is similar to the the stability function (4.3) of the implicit method for any finite nonzero z, $\widetilde{\mathbf{M}}(z) \sim \widehat{\mathbf{M}}(z)$. Therefore its spectral radius is uniformly bounded below one for all $z = h\mu$ of interest. By standard numerical ODE arguments [11] the equation (A.7) implies convergence of global errors to zero at a rate $||e_n|| = \mathcal{O}(h^{\min(p,q)})$. \square

The discussion can be extended to IMEX-TSRK method applied to semi-discrete partial differential equations. It is well known that Runge-Kutta methods can suffer severe order reductions in the presence of non-homogeneous boundary conditions or nonzero source terms [15]. Consider the system

$$y' = \mathbf{V} y + b(t) \,,$$

where **V** is a spatially discretized differential operator and b(t) represents the non-homogeneous boundary or source terms. This system can be cast in the PR form (A.1) after a transformation of variables that diagonalizes **V**, and after identifying $-\mu\phi(t) + \phi'(t)$ with the transformed b(t). The analysis reveals that the IMEX-TSRK method applied to this linear PDE system converges with order min(p, q).

Appendix B. Coefficients of the third order IMEX TSRK method. The parameters are found using the Differential Evolution ga optimization package by maximizing the length of the imaginary axis segment included in the stability region of the explicit component. The stiff accuracy condition (5.13b) and $A(\alpha)$ stability of the implicit component are included as optimization constraints. For good quality of solutions the genetic optimizer ga is run multiple times; each run is initialized with the previous result.

$$\mathbf{c} = \begin{bmatrix} 0 & 1 \end{bmatrix}^{T} \\ \mathbf{u} = \begin{bmatrix} 1 & 0 \end{bmatrix}^{T} \\ \mathbf{A} = \begin{bmatrix} 0 & 0 \\ 1.825912381819746 & 0 \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & -0.325912381819746 \end{bmatrix} \\ \hat{\mathbf{A}} = \begin{bmatrix} 5/12 & 0 \\ 0.815364469611720 & 5/12 \end{bmatrix} \\ \hat{\mathbf{B}} = \begin{bmatrix} 1/2 & 1/12 \\ -1/12 & -0.148697802945053 \end{bmatrix} \\ \vartheta = 0 \\ \mathbf{v} = \begin{bmatrix} 0.81536446961172 & 5/12 \end{bmatrix}^{T} \\ \mathbf{w} = \begin{bmatrix} -1/12 & -0.148697802945053 \end{bmatrix}^{T}$$

Appendix C. Coefficients of the fourth order IMEX TSRK method. The parameters are found using the optimization procedure explained in Appendix B.

Appendix D. Another fourth order IMEX TSRK method. The implicit component method is L-stable and is taken from [1].

$$\mathbf{c} = \begin{bmatrix} -0.19320190561126 & -0.58689424506961 & 1.08752332811466 \end{bmatrix}^T$$

$$\mathbf{u} = \begin{bmatrix} 0.45705571481934 & 1.05195992030028 & 0.15144080311463 \end{bmatrix}^T$$

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.130476793083096 & 0 & 0 \\ 1.649241112842109 & 1.814778592781876 & 0 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 0.39936246636454 & -0.16633596050061 & 0.03082730334415 \\ 0.51702376261274 & -0.18175387306707 & -0.00068100739809 \\ -5.84960861008881 & 3.22359516594067 & 0.40095792975345 \end{bmatrix}$$

$$\widehat{\mathbf{A}} = \begin{bmatrix} 0.5 & 0 & 0 \\ 0.55515820921130 & 0.5 & 0 \\ -0.27897090290997 & 2.32682280748097 & 0.5 \end{bmatrix}$$

$$\widehat{\mathbf{B}} = \begin{bmatrix} 0.01138595046334 & 0.04659103146040 & -0.29412317271565 \\ -0.48129318880262 & 0.30924798197004 & -0.41804732714804 \\ -2.38622282079758 & 0.99017411095761 & 0.08716093649826 \end{bmatrix}$$

$$\vartheta = 0$$

$$\mathbf{v} = \begin{bmatrix} -0.70240474564317 & 2.11852316846112 & 0.39319598421807 \end{bmatrix}^T$$

$$\mathbf{w} = \begin{bmatrix} -2.07554769770216 & 0.84049470544433 & 0.42573858522182 \end{bmatrix}^T$$

Appendix E. Coefficients of the sixth order IMEX TSRK method.

The development of this pair starts with a predetermined L-stable implicit component from [1]. The only free parameters for the explicit component the entries of the A matrix. The coefficients B of the nonstiff method result from the explicit stage order conditions (2.3) and the internal consistency conditions (2.11). The A coefficients are computed using a numerical optimization procedure where we optimize the explicit stability region such as to contain the longest possible interval along the imaginary axis. We have also considered maximizing the area of a half ellipse included in the stability region, and with one semi-axis overlapping the imaginary axis; the results are similar to the ones discussed below and are not reported here. The optimization process is done in two steps. First, we explore the parameter space using the genetic algorithm function ga in MATLAB optimization toolbox. The best member of this process is then taken as the starting point for MATLAB's fminsearch routine, which locally refines the solution and provides a sufficient number of accurate digits.

```
\mathbf{c} = \begin{bmatrix} -0.40455452705961 & -0.26488149320550 & 0.05730060498812 & 0.35370097422467 & 0.48881518147020 \end{bmatrix}^T
\mathbf{u} = \begin{bmatrix} 0.0002157372318872 & 0.0001354655498456 & 0.0000469196256648 & 0.0000256681792066 & 0.0000230139042425 \end{bmatrix}^T
               0
                                       0
      -0.411067755933170
                                       0
                                                                                               0
                                                                                               0
      -2.184767292491065
                              0.988337848532673
      -1.933520824847556
                             -0.039861058310592
                                                   0.757541697266156
      -0.605970379043349
                             -1.561665570720364
                                                   1.243988457457954 0.094790995264783
                                                                                          -4.93339342096474
      -1.28995988436797
                           3.01509152040882
                                                -4.63540469340360 7.43932768849976
                                                                                          -3.07365595840746
      -1.07127144661267
                           2.48892090515797
                                                -3.73103986608896 5.53336809422864
\mathbf{B} = \begin{bmatrix} -0.24158168198390 \end{bmatrix}
                           0.58513684309224
                                               -0.97567738262348 1.29786646201535
                                                                                            0.58803272807197
                                                                                          -0.64603052144256\\
      -0.67042428681047
                           1.66052413771998
                                               -2.85719727361944 4.08269477244835
      -1.51656343997187
                           3.67738907820218
                                               -5.98650039451221 8.41545308430861
                                                                                          -3.27208363610734
                                     0
              0.5
                                                           0
                                                                                                                             (E.1)
                                    0.5
      -0.21971694115244
                                                           0
                                                                                             0
      0.34677391973940
                            -0.87844580948604
                                                          0.5
                                                                                             0
                                                                               0.5
       0.06601787269656
                            -1.14976271542161
                                                  -0.08264051407691
      -1.69420405359670
                            0.16618320114019
                                                  -0.74229361529808
                                                                       0.20048354140037
                                                                                            0.5
                                                                                                -6.35888623613472
       -1.42651400380231
                             3.34568800074286
                                                   -5.21803000182508
                                                                          8.75340345119152
      -1.85213037155704
                             4.34496259959515
                                                   -6.76681774290671
                                                                         11.10081457512045
                                                                                               -7.37185814675508
\hat{\mathbf{B}} =
      -4.26388543763339
                             9.73682010462664
                                                  -13.73465558507126
                                                                         18.29003094666633
                                                                                               -9.93929061422790
      -10.19622755259608
                            22.71041154118940 -29.47781369898124
                                                                         33.88465664359473
                                                                                               -15.90091493400099
     -16.69367618197397
                            36.82959193421020
                                                 -46.44024346854481
                                                                         51.34375399377619
                                                                                               -22.98075715573894
\vartheta = 0
     \begin{bmatrix} 1.38310762038221 & -1.20955981252655 & 0.11281066473162 & -1.91985715272309 & 2.64595762329489 \end{bmatrix}^T
\mathbf{w} = \begin{bmatrix} -4.70654020323220 & 10.88079796245341 & -15.26460646566261 & 18.34573036115907 & -9.26784059787675 \end{bmatrix}^T
```

7

Appendix F. IMEX TSRK methods on Van der Pol equation. We consider the van der Pol equation

$$\frac{d}{dt} \begin{bmatrix} y \\ z \end{bmatrix} = f(y,z) + g(y,z) = \begin{bmatrix} z \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ ((1-y^2)z - y)/\epsilon \end{bmatrix}, \quad 0 \le t \le 0.55139, \quad (\text{F.1})$$

with parameter values taken from [2]

$$\epsilon = 10^{-6}, \quad y(0) = 2, \quad z(0) = -\frac{2}{3} + \frac{10}{81}\epsilon - \frac{292}{2187}\epsilon^2 - \frac{1814}{19683}\epsilon^3 + \mathcal{O}(\epsilon^4).$$
(F.2)

We integrate equation (F.1) treating the first term on the right explicitly and the other term implicitly with different fixed step sizes. The results at the final time are compared against the MATLAB reference solution. The errors for the stiff component z are plotted against step sizes in Figure F.1. All IMEX-RK methods except ARS(2,2,2) exhibit order reduction. All IMEX-BDF methods display the expected orders. IMEX-TSRK(2,3) and IMEX-TSRK(3,4) also show satisfactory convergence behavior and still retain the advantage of higher accuracy compared to the corresponding IMEX-BDF methods of the same order. IMEX-TSRK(5,6) suffers from order reduction in the high accuracy regime. This is also the case for another high order method KC785. The IMEX-RK methods use more right hand side evaluations per step than the IMEX-TSRK methods, and exhibit order reduction. This test case illustrates the benefit of the proposed IMEX-TSRK schemes.

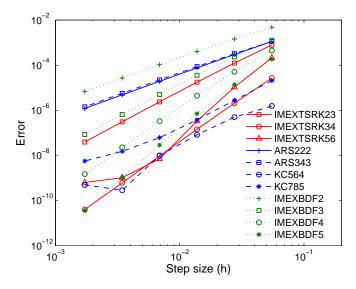
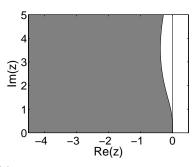
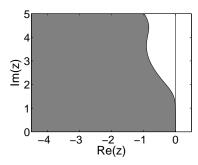


FIG. F.1. Convergence of IMEX-TSRK, IMEX-RK, and IMEX-BDF schemes for the stiff variable z of the van der Pol equation (F.1) with $\epsilon = 10^{-6}$. The solution errors (measured at the final time) are plotted against the simulation time step h.

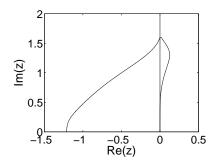
Appendix G. Stability regions for the proposed IMEX-TSRK methods of order three and four. The stability regions are shown in Figure G.1. The regions of combined stability for all choices of the stiff stability region \mathcal{S} include an open subset of the left complex half plane (thus allowing conditional stability for positive step sizes), as well as a nontrivial part of the imaginary axis (which is desirable for certain PDEs such as advection equations where the Jacobian eigenvalues after spatial discretization could have large imaginary parts [15]).

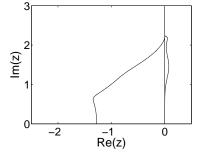




ity region.

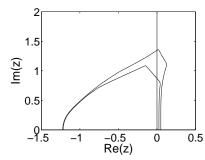
(a) The third order method. Implicit stabil- (b) The fourth order method. Implicit stability region.

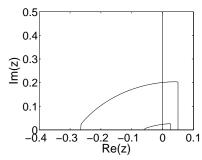




ity region.

(c) The third order method. Explicit stabil- (d) The fourth order method. Explicit stability region.



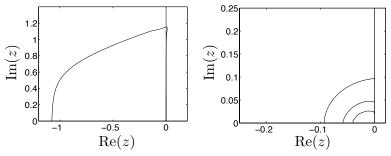


(e) The third order method. Constrained (f) The fourth order method. Constrained explicit stability regions \mathcal{N}_{α} for $\alpha \in \text{explicit}$ stability regions \mathcal{N}_{α} for α $\{60^{\circ}, 75^{\circ}\}.$ $\{60^{\circ}, 75^{\circ}\}.$

Fig. G.1. (a),(b) Stability regions for the implicit parts of the proposed IMEX-TSRK methods. (c),(d) Stability regions for the explicit parts. (e),(f) Explicit stability regions \mathcal{N}_{α} in (4.4) are constrained by the $A(\alpha)$ stability of the implicit part. The explicit stability regions \mathcal{N}_{α} shown here correspond to $\alpha = 60^{\circ}$ (outer contours) and $\alpha = 75^{\circ}$ (inner contours).

Appendix H. Stability regions for the proposed IMEX-TSRK method of order six.

The stability regions are shown in Figure H.1. The regions of combined stability for different choices of the stiff stability region \mathcal{S} are smaller than the unconstrained stability region (i.e., the standard stability region of the explicit method). Nevertheless, they include an open subset of the left complex half plane (thus allowing conditional stability for positive step sizes), as well as a part of the imaginary axis.



(a) The sixth order method. Unconstrained explicit stability region. Strained explicit stability regions \mathcal{N}_{α} for $\alpha \in \{45^{\circ}, 75^{\circ}, 90^{\circ}\}.$

Fig. H.1. (a) Stability regions for the explicit parts of the proposed IMEX-TSRK methods. (b) The explicit stability regions \mathcal{N}_{α} in (4.4) are constrained by the $A(\alpha)$ stability of the implicit part. The explicit stability regions \mathcal{N}_{α} shown here correspond to $\alpha=45^{\circ}$ (outer contours), $\alpha=75^{\circ}$ (middle contours), and $\alpha=90^{\circ}$ (inner contours).